

Received August 26, 2021, accepted September 2, 2021. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2021.3113080

Saturated Output-Feedback Hybrid Reinforcement Learning Controller for Submersible Vehicles Guaranteeing Output Constraints

OMID ELHAKI¹, KHOSHNAM SHOJAEI^{1,2}, DECLAN SHANAHAN³,
AND ALLAHYAR MONTAZERI^{1,3}

¹Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Najafabad 85141-43131, Iran

²Digital Processing and Machine Vision Research Center, Najafabad Branch, Islamic Azad University, Najafabad 85141-43131, Iran

³Department of Engineering, Lancaster University, Lancaster, England LA1 4YW, U.K.

Corresponding author: Allahyar Montazeri (a.montazeri@lancaster.ac.uk)

This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/V027379/1 and Grant EP/R02572X/1, and in part by the National Centre for Nuclear Robotics (NCNR).

ABSTRACT In this brief, we propose a new neuro-fuzzy reinforcement learning-based control (NFRLC) structure that allows autonomous underwater vehicles (AUVs) to follow a desired trajectory in large-scale complex environments precisely. The accurate tracking control problem is solved by a unique online NFRLC method designed based on actor-critic (AC) structure. Integrating the NFRLC framework including an adaptive multilayer neural network (MNN) and interval type-2 fuzzy neural network (IT2FNN) with a high-gain observer (HGO), a robust smart observer-based system is set up to estimate the velocities of the AUVs, unknown dynamic parameters containing unmodeled dynamics, nonlinearities, uncertainties and external disturbances. By employing a saturation function in the design procedure and transforming the input limitations into input saturation nonlinearities, the risk of the actuator saturation is effectively reduced together with nonlinear input saturation compensation by the NFRLC strategy. A predefined funnel-shaped performance function is designed to attain certain prescribed output performance. Finally, stability study reveals that the entire closed-loop system signals are semi-globally uniformly ultimately bounded (SGUUB) and can provide prescribed convergence rate for the tracking errors so that the tracking errors approach to the origin evolving inside the funnel-shaped performance bound at the prescribed time.

INDEX TERMS Saturation function, reinforcement learning, prescribed performance, high-gain observer, interval type-2 fuzzy neural networks, multilayer neural networks.

I. INTRODUCTION

Robotics and autonomous system have found many applications in hazardous environments, especially when the area is inaccessible to the human or will impose tremendous danger to the operator to work on site. For example, significant investment has been made to develop next generation of dexterous manipulators for nuclear decommissioning applications [1], [2]. The use of quadcopter along with the manipulator helps to improve the situational awareness and autonomy of the manipulator when it interacts with

the surrounding environment without human intervention [3], [4]. Another useful robot to operate in nuclear applications is autonomous underwater robots. They are predominantly used for inspection and monitoring of the nuclear wastes stored underwater inside the nuclear ponds. Currently, in most of the applications, a specific category of underwater robots, i.e. submersible remotely operated vehicles (ROVs) are utilised. The global offshore ROV industry is predicted to worth up to \$3.5Bn with modern applications including civil and inshore inspections (CIIs) and decommissioning of oil and gas pipelines, nuclear storage facilities and accident sites, liquid storage tanks and tunnels [5]. Autonomous operation of such underwater vehicles introduces many challenges due to

The associate editor coordinating the review of this manuscript and approving it for publication was Shun-Feng Su.

the nonlinear and uncertain behavior of the vehicle. For example, this includes but not limited to finite scape time, multiple equilibrium points, limit cycles, appearing new frequencies, bifurcation, chaos, unmodeled dynamics, uncertainties, etc. These problems have attracted a lot of interest in many real world robotic and mechanical systems such as AUVs, wheeled mobile robots, robot manipulators and aerial vehicles. Classical control approaches have not guaranteed tolerable results in tough operations of uncertain nonlinear dynamic systems. To this end, it is indispensable for automatic applications to design model-free intelligent controllers with strong robustness against uncertain nonlinearities to attain a high precision tracking performance facing complex system features, which are sustainable substitutions under these conditions. To reach appropriate model-free control aims, many interesting methods have been expanded for uncertain nonlinear systems via artificial neural networks (ANNs) and fuzzy logic systems (FLSs). For example, Elhaki & Shojaei have developed an output-feedback multi-layer neural network (MNN)-based controller for AUVs in the situation of complex nonlinearities and unmeasured states [6]. In [7], an adaptive IT2FNN was employed to estimate system's nonlinear functions. The fuzzy set theory was firstly presented by Zadeh [8]. Since then, these systems are classified as type-1 fuzzy logic systems (T1FLSs) and they are widely used in many control applications. However, since T1FLSs are exact sets, they cannot take into account the uncertainty in the membership functions (MFs). Therefore, Zadeh established the idea of type-2 fuzzy logic systems (T2FLSs) in 1975 [9]. As a result, uncertainties have been more sufficiently handled by T2FLSs in control methods since the degree of MFs are themselves based on the fuzzy logic. On the other side, since T2FLS was computationally expensive, interval type-2 fuzzy logic systems (IT2FLSs) have been introduced by Liang and Mendel [10]. In [11], an IT2FLS was employed for compensating uncertainties in wireless sensor networks. But, the knowledge utilized to build rules in a FLS may be uncertain and inaccurate, which may result in the fact that rules have uncertain consequents/antecedents revealed in MFs of consequents/antecedents. To overcome this defect, many researches have fused IT2FLS and ANNs to create IT2FNNs. The IT2FNNs take the advantages of both interval type-2 fuzzy reasoning to prevail severe nonlinearities and the ability of ANNs to learn and discover the best action from the process [12]–[14]. The structure of IT2FNNs can enhance the accuracy of control systems and they are able to prevail over the limitations of classical controllers, which are parametric and structural uncertainties along with time-variant unmodeled dynamics that can lead to undesirable effects on the control systems.

Recently, a fuzzy optimal control scheme has been a topic of the research interest. In [15], the fuzzy optimal tracking control of hypersonic flight vehicles was addressed. The issue of the accelerated adaptive fuzzy optimal control of three coupled fractional-order chaotic electromechanical transducers was addressed in [16]. Reference [17] has proposed an

adaptive fuzzy inverse optimal output feedback controller for vehicular active suspension systems. In [18], an optimal fuzzy adaptive robust controller for a lower limb exoskeleton robot system has been suggested. The adaptive fuzzy finite time control problem for a class of switched nonlinear systems has been discussed in [19], and in [20], a fuzzy swing up controller and the optimal state feedback stabilization for an inverted pendulum was debated. Lately, multi-layer neural network reinforcement learning (MNNRL) has been a topic of research interest for IT2FNNs and ANNs to improve the performance of the network [21]. The MNNRL is a combination of multi-layer neural networks (MNNs) and reinforcement learning (RL) to solve the problem of dimensional explosion in case the amount of state space increases that restricts the performance of RL [22]. The RL is an adaptive intelligent branch of machine learning. The philosophy behind RL is that the closed-loop system interacts with an approximator by determining three essential signals: the system's state signals, which permits the controller to affect the system, the action signal, which is the output of the actor agent and influences the system, and the reinforcement signal, which is the output of the critic agent according to the evaluation of the actor's performance that is an estimation of a cost function [23]. In [24], a RL-based controller has been developed for a flexible two-link manipulator, and in [25], an integral RL controller with input saturation for nonlinear systems was studied. However, these methods could not deal with the dimensionality problem. In contrast, references [26] and [27] have developed deep RL-based controllers for wheeled mobile robots, but these algorithms cannot guarantee the stability of the system and suffer from the input saturation. Reference [28] has proposed an adaptive RL controller for hypersonic vehicles, but saturation functions along with the compensation of the actuators saturation nonlinearity are not taken into account and the suggested controller needs the measurements of all system's state derivatives. Here, from our opinion, some fundamental questions have arisen which are inspiring for this research work: How to design a saturated NFRLC-based method whose stability can be guaranteed via a strong Lyapunov stability analysis? Can we design a hybrid NFRLC-estimation-based framework including a MNN, which acts as the actor agent and can deal with the dimensionality problem, and an IT2FNN, which acts as the critic agent and can handle uncertain nonlinearities strongly in order to take the advantages of MNNs and IT2FNNs simultaneously? Finding answers for the aforesaid questions is one of the main contributions of this research paper.

Since unmeasured states may exist in some control structures due to the lack of sensors or the implementation cost, state feedback controllers may not be applied and a state observer should be employed in the control system. In this respect, some influential observer-based controllers have been developed such as dynamic event-triggered observer-based PID controller [29], cascade predictive observer [30], extended state observer [31], and

high-gain observer (HGO) [32]. Yet, the most of cited control methods were designed to ensure the stability of the closed-loop control system by using Lyapunov stability principle that may fail to realize the desired tracking performance. In this context, the prescribed performance control (PPC) [33]–[37] is a good option to deal with this problem. This technique provides a nonlinear transformation from an unconstrained tracking problem to a prescribed constrained tracking problem, which is valid to cope with prescribed performance tracking problems and the tracking errors are also guaranteed to be within some prescribed performance bounds (PBs). In the last few years, many control methodologies have employed PPC method to improve the performance of the control systems. In [38], a PPC-based controller is designed for uncertain nonlinear systems with unknown control directions. A new data-driven model-free adaptive terminal sliding mode control for a class of discrete nonlinear systems with a prescribed performance has been designed in [39]. Reference [40] has developed a fuzzy wavelet neural controller with an improved prescribed performance for a micro-electromechanical system gyroscope. The problem of the adaptive neural output-feedback tracking control for a class of switched uncertain nonlinear systems with a prescribed performance is addressed in [41], and in [42], the problem of the finite-time prescribed performance adaptive fuzzy control for a class of unknown nonlinear systems was investigated.

The control strategy in this paper is to design a multi-objective model-free intelligent controller including AC-NFRLC, MNNs, IT2FNNs, saturation functions, PPC, and a state observer with a strong Lyapunov-based stability proof. Chiefly, main contributions of this study are listed as follows: (1) A novel hybrid framework for AC-NFRLC mechanism based on MNNs and IT2FNNs is designed. (2) The substructure of the proposed AC-NFRLC system inherits the advantages of MNNs, which is dealing with the dimensionality problem, and IT2FNNs, which are fuzzy reasoning and learning from the process. This fact leads to a strong AC-NFRLC structure for compensating uncertain nonlinearities. (3) The adaptive laws for the learning process are derived by the Lyapunov theory. (4) To deal with unmeasured states, a HGO is efficiently employed. (5) An adaptive robust controller is effectively fused with the proposed AC-NFRLC framework to compensate the effects of external disturbances and function approximation errors. (6) An efficient saturation function is utilized in the design procedure to bound the error signals and to reduce the actuator saturation risk. (7) The risk of actuator saturation is diminished by learning and compensating the actuator saturation nonlinearity via the proposed AC-NFRLC method. (8) To achieve a desired tracking performance with predefined transient and steady-state response characteristics, the PPC technique is employed. (9) The stability of the entire closed-loop control system is strongly proved by the Lyapunov theory. The integration of items (1)–(9) brings many challenges in designing the proposed controller and deriving mathematical equations. In this paper,

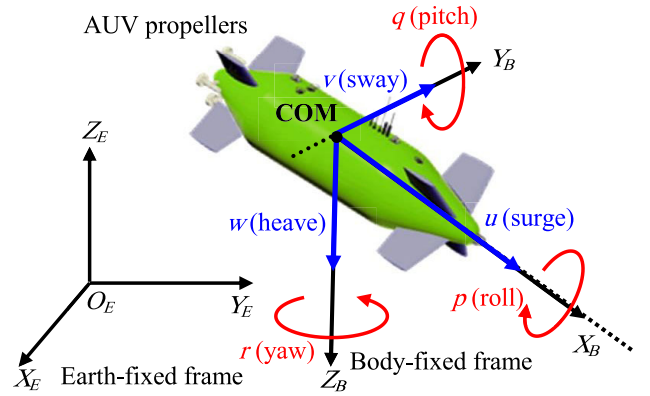


FIGURE 1. Three-dimensional configuration of a fully-actuated AUV.

we successfully coped with the control objectives presented in (1)–(9) along with a strong Lyapunov-based stability proof. To the best of authors' knowledge, it is the first time that items (1)–(9) are considered in the context of submersible vehicles, leading to designing a novel saturated intelligent controller with the capability of handling external disturbances, dimensionality, and unknown nonlinearities in real-time with a prescribed performance without velocity measurements.

The arrangement of this paper is given by the following steps. In Section II, some preliminaries are given. The framework and design ideology of IT2FNNs along with MNNs and the design process of the proposed method are given in Section III. In Section IV, simulation results are provided to show the controller efficacy. Finally, the concluding statements are given in Section V.

II. PRELIMINARIES

A. NOTATIONS

In this paper, the following notations are used otherwise they will be indicated. $\text{diag}[\bullet]$ is a diagonal matrix, $0_{n \times n}$ indicates a $n \times n$ matrix of zeros, I_n denotes a $n \times n$ identity matrix, $\|\bullet\|_F$ stands for the Frobenius norm, $\|\bullet\|$ shows the Euclidean norm, $\text{blkdiag}[\bullet]$ represents a block diagonal matrix, and $\text{col}(A, B)$ is a column of a partitioned matrix with arbitrary sub-matrices A and B where A and B are placed in the top and bottom blocks, respectively.

B. SYSTEM DESCRIPTION

Consider the following kinematic and dynamic motion equations of AUVs [43]:

$$\begin{aligned} \dot{q}_1 &= J_1(q_2)v_1, \\ M_1\dot{v}_1 &= -C_1(v_1)v_2 - D_1v_1 - u_{d1}(v_1) + \tau_{s1} + \delta_1, \\ \dot{q}_2 &= J_2(q_2)v_2, \\ M_2\dot{v}_2 &= -C_1(v_1)v_1 - C_2(v_2)v_2 - D_2v_2 - u_{d2}(v_2) \\ &\quad - g_1(q_2) + \tau_{s2} + \delta_2, \end{aligned} \quad (1)$$

where $q_1 = [x, y, z]^T$ denotes the vector of AUV's position, $q_2 = [\phi, \theta, \psi]^T$ represents the orientation of the AUV, $v_1 = [u, v, w]^T$ and $v_2 = [p, q, r]^T$ are the velocity vectors,

$\delta_1 = [\delta_x, \delta_y, \delta_z]^T$ and $\delta_2 = [\delta_\phi, \delta_\theta, \delta_\psi]^T$ denote the bounded external disturbance forces and torques. Eq. (1) is rewritten as

$$\begin{cases} \dot{q} = J_t(q)v, \\ M_t \dot{v} = -C_t(v)v - D_t v - g(q) - u_d(v) + \tau_s + \delta, \end{cases} \quad (2)$$

where $q = [q_1^T, q_2^T]^T$ is the vector of positions and orientations, $J_t(q) = \text{blkdiag}[J_1, J_2] \in \mathbb{R}^{6 \times 6}$ specifies a rotational matrix, $v = [v_1^T, v_2^T]^T \in \mathbb{R}^6$ is the velocity vector, $M_t = \text{blkdiag}[M_1, M_2] \in \mathbb{R}^{6 \times 6}$ stands for the inertia matrix that is positive-definite and symmetric, $C_t = [\text{col}(0_{3 \times 3}, C_1), \text{col}(C_1, C_2)] \in \mathbb{R}^{6 \times 6}$ denotes the matrix of Coriolis and centripetal forces which is skew-symmetric, $D_t = \text{blkdiag}[D_1, D_2] \in \mathbb{R}^{6 \times 6}$ shows the damping forces matrix, $g(q) = [0, 0, 0, g_1^T(q_2)]^T \in \mathbb{R}^6$ is the gravitational forces and buoyancy moments, $u_d = [u_{d1}^T, u_{d2}^T]^T \in \mathbb{R}^6$ is the vector of unmodeled dynamics and nonlinearities/uncertainties, $\tau_s = [\tau_{us}, \tau_{vs}, \tau_{ws}, \tau_{ps}, \tau_{qs}, \tau_{rs}]^T \in \mathbb{R}^6$ denotes the saturated input control forces and torques, and $\delta = [\delta_1^T, \delta_2^T]^T \in \mathbb{R}^6$ stands for the vector of bounded environmental disturbances. For more details on AUVs dynamics, readers are referred to [43]. The saturated input forces and torques are

$$\tau_s = \begin{cases} \tau_{\max}, & r\tau \geq \tau_{\max} \\ r\tau, & -\tau_{\max} < r\tau < \tau_{\max} \\ -\tau_{\max}, & r\tau \leq -\tau_{\max} \end{cases} \quad (3)$$

where τ_s , τ , and τ_{\max} denote the vectors of saturated inputs, non-saturated inputs, and the limit of the actuators, respectively, and r is a ratio between τ and τ_s . The input saturation nonlinearity $d_\tau(\tau) = \tau_s - \tau$, that is a NLIP term and will be compensated in the control design procedure, is given by

$$d_\tau(\tau) = \begin{cases} \tau_{\max} - \tau, & r\tau \geq \tau_{\max} \\ (r-1)\tau, & -\tau_{\max} < r\tau < \tau_{\max} \\ -\tau_{\max} - \tau, & r\tau \leq -\tau_{\max} \end{cases} \quad (4)$$

C. PRESCRIBED PERFORMANCE TRANSFORMATION

To obtain PPC objectives, errors (i.e. $e = q - q_d$) should evolve inside a funnel-shaped set during the time [44], [45]. A smooth limited function $\eta_i : \mathbb{R}^+ \rightarrow \mathbb{R}^+, i = 1, \dots, 6$, is a PB, if η_i is decreasing, and $\lim_{t \rightarrow \infty} \eta_i(t) = \eta_{i\infty}$ [33]. The PPC is provided if $\eta_{li}(t) \leq e_i(t) \leq \eta_{ui}(t)$ be true where e_i is i th element of e , η_{li} and η_{ui} are the pre-set limits of e_i , and η_i is defined as $\eta_i(t) := (\eta_{i0} - \eta_{i\infty})\exp(-a_i t) + \eta_{i\infty}$ where $a_i \in \mathbb{R}^+$ is a lower bound on the convergence rate, $\eta_{i0} > \eta_{i\infty} \in \mathbb{R}^+$ are design factors, $\eta_{i\infty}$ is an upper bound for the final steady-state error. The PBs can be set as $\eta_{li} = -\alpha_i \eta_i$, $\eta_{ui} = \beta_i \eta_i$ where $\alpha_i, \beta_i \in \mathbb{R}^+$ are tunable factors.

Assumption 1: The errors meet $\eta_{li}(0) \leq e_i(0) \leq \eta_{ui}(0)$.

The next nonlinear transformation is used here:

$$\epsilon_e = \rho^{-1}(e) = [\rho_1^{-1}(e_1), \dots, \rho_6^{-1}(e_6)]^T. \quad (5)$$

which is invertible as $e = \rho(\epsilon_e)$.

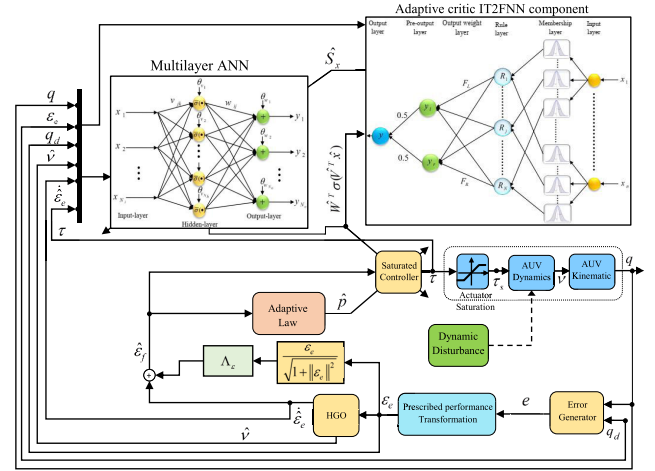


FIGURE 2. Representation of the proposed intelligent system.

Fact 1. Transformation (5) should meet the followings:

$$\begin{cases} \lim_{\epsilon_{ei} \rightarrow +\infty} \rho_i(\epsilon_{ei}) = \eta_{ui}, \\ \lim_{\epsilon_{ei} \rightarrow -\infty} \rho_i(\epsilon_{ei}) = \eta_{li}, \end{cases} \quad (6)$$

where Fact 1 implies that tracking errors will settle inside the PBs and converge to a neighbourhood of the zero if Assumption 1 holds true and $\epsilon_e \in \mathcal{L}_\infty$ is ensured by a suitable controller. A candidate for transformation (5) that meets the properties of Fact 1 is

$$e_i = \rho_i(\epsilon_{ei}) = \frac{\eta_{ui} - \eta_{li}}{\pi} \arctan(\epsilon_{ei}) + \frac{\eta_{ui} + \eta_{li}}{2}. \quad (7)$$

From (7), the transformed errors $\epsilon_{ei}(t)$ are

$$\epsilon_{ei} = \rho_i^{-1}(e_i) = \tan\left(\frac{\pi}{2} \times \frac{2e_i - \eta_{ui} - \eta_{li}}{\eta_{ui} - \eta_{li}}\right). \quad (8)$$

Time derivation of (8) yields

$$\dot{\epsilon}_{ei} = \frac{\partial \epsilon_{ei}}{\partial e_i} \dot{e}_i + \Psi_i, \quad (9)$$

where $\Psi_i = \frac{\partial \epsilon_{ei}}{\partial \eta_{ui}} \dot{\eta}_{ui} + \frac{\partial \epsilon_{ei}}{\partial \eta_{li}} \dot{\eta}_{li}$. Utilizing (7) and (8) gives

$$\frac{\partial \epsilon_{ei}}{\partial e_i} = \frac{\pi}{\eta_{ui} - \eta_{li}} \sec^2\left(\frac{\pi}{2} \times \frac{2e_i - \eta_{ui} - \eta_{li}}{\eta_{ui} - \eta_{li}}\right) > 0, \quad (10)$$

$$\frac{\partial \epsilon_{ei}}{\partial \epsilon_{ei}} = \frac{\eta_{ui} - \eta_{li}}{\pi(1 + \epsilon_{ei}^2)} > 0, \quad (11)$$

which show a strictly increasing relation between ϵ_{ei} and e_i . Finally, using $\dot{e} = \dot{q} - \dot{q}_d$, (2), and (9) gives

$$\dot{\epsilon}_e = \mathcal{D}v + \kappa, \quad (12)$$

where $\mathcal{D} = TJ_t$, $\kappa(\epsilon_e, \eta_u, \eta_l, \dot{\eta}_u, \dot{\eta}_l, q_d, \dot{q}_d) = \Psi - T\dot{q}_d$, $T = \text{diag}[\partial \epsilon_{e1}/\partial e_1, \dots, \partial \epsilon_{e6}/\partial e_6]$, $\Psi = [\Psi_1, \dots, \Psi_6]^T$, $\eta_u = [\eta_{u1}, \dots, \eta_{u6}]^T$, $\eta_l = [\eta_{l1}, \dots, \eta_{l6}]^T$, $\dot{\eta}_u = [\dot{\eta}_{u1}, \dots, \dot{\eta}_{u6}]^T$, and $\dot{\eta}_l = [\dot{\eta}_{l1}, \dots, \dot{\eta}_{l6}]^T$.

III. NFRLC METHODOLOGY DESIGN

In this section, a neuro-fuzzy reinforcement learning-based controller is designed for the accurate tracking of AUVs in which the actor agent is a multilayer neural network that can deal with the dimensionality problem, and the critic agent is an interval type-2 fuzzy neural network that can handle uncertainties strongly. By calculating velocity and acceleration from (12) and its derivative as $v = \mathcal{D}^{-1}\dot{\epsilon}_e - \mathcal{D}^{-1}\kappa$, $\dot{v} = \mathcal{D}^{-1}\ddot{\epsilon}_e - \mathcal{D}^{-1}\dot{\mathcal{D}}\mathcal{D}^{-1}\dot{\epsilon}_e + \mathcal{D}^{-1}\dot{\mathcal{D}}\mathcal{D}^{-1}\kappa - \mathcal{D}^{-1}\dot{\kappa}$, substituting them into the second equation of (2), and the product of \mathcal{D}^{-T} in the resultant equation, one gets the following second-order error model in terms of unconstrained errors:

$$M(\epsilon_e)\ddot{\epsilon}_e + C(\epsilon_e, \dot{\epsilon}_e)\dot{\epsilon}_e + D(\epsilon_e)\dot{\epsilon}_e - \zeta = \mathcal{D}^{-T}\tau + \delta_p, \quad (13)$$

where $M(\epsilon_e) = \mathcal{D}^{-T}M_t\mathcal{D}^{-1}$, $C(\epsilon_e, \dot{\epsilon}_e) = \mathcal{D}^{-T}(C_t - M_t\mathcal{D}^{-1}\dot{\mathcal{D}})\mathcal{D}^{-1}$, $D(\epsilon_e) = \mathcal{D}^{-T}D_t\mathcal{D}^{-1}$, $\delta_p = \mathcal{D}^{-T}\delta$ where $|\delta_{p_i}| \leq B_{\delta_i}$, $i = 1, \dots, 6$, and NLIP uncertain term ζ is

$$\zeta = M(\epsilon_e)\dot{\kappa} + C(\epsilon_e, \dot{\epsilon}_e)\kappa + \mathcal{D}^{-T}C_t(v)\mathcal{D}^{-1}\kappa + D(\epsilon_e)\kappa - \mathcal{D}^{-T}g(q) - \mathcal{D}^{-T}u_d(v) + \mathcal{D}^{-T}d_\tau(\tau). \quad (14)$$

Property 1: Because \mathcal{D} is a full-rank matrix by recalling (10), the following items are valid for (14).

P1.1: $M(\epsilon_e) = M^T(\epsilon_e) > 0$, $\lambda_m\|x\|^2 \leq x^T Mx \leq \lambda_M\|x\|^2$, $\forall x \in \mathbb{R}^6$ where $0 < \lambda_m < \lambda_M < \infty$, and $\lambda_m := \min_{\forall \epsilon_e \in \mathbb{R}^6} \lambda_{\min}(M(\epsilon_e))$, $\lambda_M := \max_{\forall \epsilon_e \in \mathbb{R}^6} \lambda_{\max}(M(\epsilon_e))$.

P1.2: $D(\epsilon_e) = D^T(\epsilon_e) > 0$ where $\lambda_d\|x\|^2 \leq x^T Dx \leq \lambda_D\|x\|^2$, $\forall x \in \mathbb{R}^6$ where $0 < \lambda_d < \lambda_D < \infty$, and $\lambda_d := \min_{\forall \epsilon_e \in \mathbb{R}^6} \lambda_{\min}(D(\epsilon_e))$, $\lambda_D := \max_{\forall \epsilon_e \in \mathbb{R}^6} \lambda_{\max}(D(\epsilon_e))$.

P1.3: The matrix $C(\epsilon_e, \dot{\epsilon}_e)$ has the following properties:

- (i) $x^T (\dot{M}(\epsilon_e) - 2C(\epsilon_e, \dot{\epsilon}_e))x = 0$, $\forall x \in \mathbb{R}^6$,
- (ii) $C(\epsilon_e, x_1)x_2 = C(\epsilon_e, x_2)x_1$,
- (iii) $C(\epsilon_e, x_1 + x_2)y = C(\epsilon_e, x_1)y + C(\epsilon_e, x_2)y$,
- (iv) $\|C(\epsilon_e, x_1)x_2\| \leq B_c\|x_1\|\|x_2\|$ where $B_c \in \mathbb{R}^+$.

Next, the following saturated filtered tracking error is introduced:

$$\varepsilon_f = \dot{\epsilon}_e + \Lambda_\varepsilon \frac{\epsilon_e}{\sqrt{1 + \|\epsilon_e\|^2}}, \quad (15)$$

where Λ_ε is a gain matrix. Employing (13), (15), items (ii) and (iii) of P1.3 gives

$$\begin{aligned} M(\epsilon_e)\dot{\varepsilon}_f &= -C(\epsilon_e, \dot{\epsilon}_e)\varepsilon_f - D(\epsilon_e)\varepsilon_f + \zeta + \gamma \\ &\quad + \mathcal{D}^{-T}\tau + \delta_p, \end{aligned} \quad (16)$$

$$\begin{aligned} \gamma &= M(\epsilon_e)\Lambda_\varepsilon\dot{\epsilon}_e(1 + \|\epsilon_e\|^2)^{-\frac{3}{2}} + D(\epsilon_e)\Lambda_\varepsilon\epsilon_e(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \\ &\quad + C(\epsilon_e, \dot{\epsilon}_e)\Lambda_\varepsilon\epsilon_e(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}}, \end{aligned} \quad (17)$$

where γ is bounded by using item (iv) of P1.3, P1.1, and P1.2 as $\|\gamma\| \leq \iota_1\|x_\varepsilon\| + \iota_2\|x_\varepsilon\|^2$ where $x_\varepsilon = [\epsilon_e^T(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}}, \varepsilon_f^T]^T$, and $\iota_1, \iota_2 \in \mathbb{R}^+$ are unknown.

From a study of [46], MNNs can solve the effect of dimensionality curse (high dimensionality of the input space) and

can be used to estimate unknown NLIP term $\zeta(x)$ as

$$\zeta_i = \sum_{j=1}^{N_h} \left[w_{ij} \bar{\sigma} \left(\sum_{k=1}^{N_i} v_{jk} x_k + \theta_{vj} \right) + \theta_{wi} \right], \quad (18)$$

where $i = 1, \dots, N_o$, N_h, N_i and N_o are the number of hidden-layer, input-layer, and output-layer cells, respectively, w_{ij} and v_{jk} are the NN weights, θ_{vj} and θ_{wi} show threshold offsets, and $\bar{\sigma}(\xi) = 1/(1 + e^{-\xi})$ is an activation function. Eq. (18) is rewritten as $\zeta = W^T \sigma(V^T x)$ where $W^T \in \mathbb{R}^{N_o \times (N_h+1)}$, $V^T \in \mathbb{R}^{N_h \times (N_i+1)}$ are the ideal ANN matrices and their first columns contain thresholds θ_{vj} and θ_{wi} , $x = [1, \epsilon_e^T, \dot{\epsilon}_e^T, v^T, \tau^T, \eta_l^T, \dot{\eta}_l^T, \ddot{\eta}_l^T, \eta_u^T, \dot{\eta}_u^T, \ddot{\eta}_u^T, q^T, \dot{q}_d^T, \ddot{q}_d^T]^T \in \mathbb{R}^{N_i+1}$, $\zeta = [\zeta_1, \dots, \zeta_{N_o}]^T \in \mathbb{R}^{N_o}$, and $\sigma(V^T x) = [1, \bar{\sigma}(V_{r_1}^T x), \dots, \bar{\sigma}(V_{r_{N_h}}^T x)]^T \in \mathbb{R}^{N_h+1}$ where $V_{r_j}^T, j = 1, \dots, N_h$ is j th row of V^T . For the continuous function $\zeta(x) : \omega_n \rightarrow \mathbb{R}^{N_o}$ where $\omega_n \subset \mathbb{R}^{N_i+1}$ denotes a compact set, there exist ideal ANN weights, thresholds, and some numbers of hidden-layer cells such that $\zeta(x) = W^{*T} \sigma(V^{*T} x) + e_x(x)$ where $e_x(x) \in \mathbb{R}^{N_o}$ is the ANN estimation error that is bounded on ω_n in a manner that $|e_{x_i}| \leq B_{x_i}$, $i = 1, \dots, N_o$, $\forall x \in \omega_n$ where $B_{x_i} \in \mathbb{R}^+$. The ANN matrices $W^{*T} \in \mathbb{R}^{N_o \times (N_h+1)}$, $V^{*T} \in \mathbb{R}^{N_h \times (N_i+1)}$ are given by

$$(W^*, V^*) := \arg \min_{(W, V)} \left\{ \sup_{x \in \omega_n} \|W^T \sigma(V^T x) - \zeta(x)\| \right\}. \quad (19)$$

Assumption 2: Matrices W^* and V^* are bounded on ω_n so that $\|W^*\|_F \leq B_w$, $\|V^*\|_F \leq B_v$ where $B_w, B_v \in \mathbb{R}^+$.

However, matrices W^* , V^* and system derivatives are generally unknown and $\zeta(x)$ is substituted by its approximation, i.e. $\hat{\zeta}(\hat{x}) = \hat{W}^T \sigma(\hat{V}^T \hat{x})$, where \hat{W} and \hat{V} denote the estimated ANN matrices. Since it is assumed that system derivatives are immeasurable, a HGO is used based on the following lemma.

Lemma 1 [47]: Suppose that the transformed error, i.e. ϵ_e , up to its $n-1$ derivatives are bounded that is $\|\epsilon_e^{(k)}\| \leq B_k$ where $B_k \in \mathbb{R}^+$, and the subsequent system is considered:

$$\begin{aligned} k_o \dot{\ell}_k &= \ell_{(k+1)}, \quad k = 1, \dots, n-1, \\ k_o \dot{\ell}_n &= -\lambda_1 \ell_n - \lambda_2 \ell_{(n-1)} - \dots - \lambda_{(n-1)} \ell_2 - \ell_1 + \epsilon_e, \end{aligned} \quad (20)$$

where $k_o \in \mathbb{R}^+$ is a design parameter. The constants λ_1 to $\lambda_{(n-1)}$ are set to allow the polynomial $\mathbb{S}^n + \lambda_1 \mathbb{S}^{n-1} + \dots + \lambda_{(n-1)} \mathbb{S} + 1$ be Hurwitz. Then, the following points are valid:

i) $\ell_{(j+1)}/k_o^j - \epsilon_e^{(j)} = -k_o \varpi^{(j+1)}$, $j = 0, 1, \dots, n-1$ where $\varpi = \ell_n + \lambda_1 \ell_{(n-1)} + \dots + \lambda_{(n-1)} \ell_1$, and $\varpi^{(j)}$ is j th derivative of ϖ . According to [47], $\ell_{(j+1)}/k_o^j$ has a tendency to $\epsilon_e^{(j)}$ asymptotically including a small error subjected to $\epsilon_e(t)$ and its j th derivatives be bounded; ii) there exist $t_1, G_j \in \mathbb{R}^+$ so that for $t > t_1$, the fact $\|\varpi^{(j)}\| \leq G_j$ is true.

Now, by employing Lemma 1, ε_f can be estimated as $\hat{\varepsilon}_f = \dot{\epsilon}_e + \Lambda_\varepsilon \epsilon_e(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} = \ell_2/k_o + \Lambda_\varepsilon \epsilon_e(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}}$ by

$$k_o \dot{\ell}_1 = \ell_2, \quad (21)$$

$$k_o \dot{\ell}_2 = -\lambda_1 \ell_2 - \ell_1 + \epsilon_e(t). \quad (22)$$

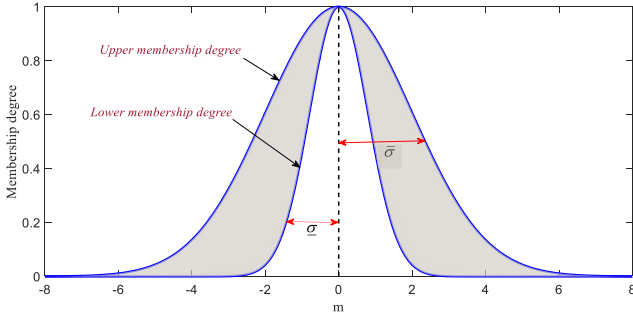


FIGURE 3. Representation of an IT2FNN MF with uncertain standard deviation.

Now, by using item (ii) of Lemma 1 and $\tilde{\varepsilon}_f = \hat{\varepsilon}_f - \varepsilon_f = \dot{\hat{\varepsilon}}_e - \dot{\varepsilon}_e$, one has $\|\tilde{\varepsilon}_f\| = \|k_o \ddot{\omega}\| \leq k_o G_2 := B_\varepsilon$ where $B_\varepsilon \in \mathbb{R}^+$. Now, the error of ANN estimation is given by [48]

$$\varsigma - \hat{\varsigma} = \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) + \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \tilde{V}^T \hat{x} + r_t + e_x(x), \quad (23)$$

where $\sigma'(\hat{V}^T \hat{x}) = [0_{N_h \times 1}, \text{diag}[\sigma'_1, \dots, \sigma'_{N_h}]]^T \in \mathbb{R}^{(N_h+1) \times N_h}$, with $\sigma'_i = d\sigma(z)/dz|_{z=\hat{V}^T \hat{x}}$, $i = 1, \dots, N_h$, $\hat{x} = [1, \epsilon_e^T, \dot{\epsilon}_e^T, \hat{v}^T, \tau^T, \eta_l^T, \dot{\eta}_l^T, \ddot{\eta}_l^T, \eta_u^T, \dot{\eta}_u^T, \ddot{\eta}_u^T, q^T, \dot{q}_d^T, \ddot{q}_d^T]^T$, and r_t in (23) is bounded as

$$\|r_t\| \leq \|W\|_F (\|\sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}\| + \|\sigma(\hat{V}^T \hat{x})\|) + \|V\|_F \|\hat{x}\| \|\hat{W}^T \sigma'(\hat{V}^T \hat{x})\|_F. \quad (24)$$

To generate an efficient reinforcement signal by collecting the error metric signal for tuning the ANN system, an IT2FNN is used here [7], [49], [50]. The major part of fuzzy systems is the fuzzy rule base that can be expressed as IF-THEN rules. Presume that the fuzzy rule base includes of N rules according to the following general rule:

R_j ($j = 1, \dots, N$): If x_1 is \tilde{A}_{j1} and \dots and x_n is \tilde{A}_{jn} then y_1 is \tilde{f}_{j1} and \dots and y_m is \tilde{f}_{jm} , where x_i ($i = 1, \dots, n$) are the inputs of the fuzzy system, y_k ($k = 1, \dots, m$) are the outputs of the fuzzy system, $\tilde{A}_{ji} = [\underline{\mu}_{ji}(x_i) \ \overline{\mu}_{ji}(x_i)]$ denotes the membership degree of j th lower and upper MF for i th input, $\tilde{f}_{jk} = [f_{jkl} \ f_{jkr}]$ is the interval consequent part of k th output. The MF of i th input is chosen as

$$\begin{cases} \underline{\mu}_{ji}(x_i) = \exp(-0.5(\frac{x_i - m_{ji}}{\underline{\sigma}_{ji}})^2), \\ \overline{\mu}_{ji}(x_i) = \exp(-0.5(\frac{x_i - m_{ji}}{\overline{\sigma}_{ji}})^2), \end{cases} \quad (25)$$

where m_{ji} is mean, $\underline{\sigma}_{ji}$ and $\overline{\sigma}_{ji}$ denote the width of j th lower and upper MF for i th input. The output of the IT2FNN is

$$y = 0.5F^T \sigma_f(x), \quad (26)$$

where $F^T = [F_L^T \ F_R^T] \in \mathbb{R}^{m \times 2N}$, $F_L = [f_{jkl}] \in \mathbb{R}^{N \times m}$ and $F_R = [f_{jkr}] \in \mathbb{R}^{N \times m}$ are the ideal weight matrices, $\sigma_f(x) = [\sigma_f^T(x) \ \overline{\sigma}_f^T(x)]^T$, $\sigma_f(x) = [\underline{\sigma}_{f1}/\sum_{j=1}^N \underline{\sigma}_{fj}, \dots, \underline{\sigma}_{fN}/\sum_{j=1}^N \underline{\sigma}_{fj}]^T$, $\overline{\sigma}_f(x) = [\overline{\sigma}_{f1}/\sum_{j=1}^N \overline{\sigma}_{fj}, \dots, \overline{\sigma}_{fN}/\sum_{j=1}^N \overline{\sigma}_{fj}]^T$, $\underline{\sigma}_{fj} = \Pi_{i=1}^n \underline{\mu}_{ji}(x_i)$, and

$\overline{\sigma}_{fj} = \Pi_{i=1}^n \overline{\mu}_{ji}(x_i)$. For a continuous function $g(x)$, an optimal matrix F exists so that

$$F := \arg \min_F \left\{ \sup_{x \in \omega_f} \|0.5F^T \sigma_f(x) - g(x)\| \right\}, \quad (27)$$

where $\omega_f \subset \mathbb{R}^n$ is a compact set.

Assumption 3: F is bounded on ω_f so that $\|F\|_F \leq B_f$, where $B_f \in \mathbb{R}^+$.

Now, the following critic-fuzzy-based cost function is considered [51]:

$$S_x = \varepsilon_f + 0.5\|\varepsilon_f\|F^T \sigma_f(x). \quad (28)$$

Since F and some states of the system are unknown in general, the following adaptive critic fuzzy component is proposed here:

$$\hat{S}_x = \hat{\varepsilon}_f + 0.5\|\hat{\varepsilon}_f\|\hat{F}^T \sigma_f(\hat{x}), \quad (29)$$

where \hat{S}_x is used later to generate the ANN weight update rules that supports the ANN to do the estimation strategy more precisely online, and corrects the system's output as well.

Remark 1: Because the adaptive critic fuzzy component in (29) can be straightly seen as a form of reinforcement signal and \hat{S}_x is more informative than the system states, an excellent control action can be resulted and a better control efficiency and performance can be obtained [52], [53].

Remark 2: The definition objective for the critic cost function (29) is to reach an optimal control action. In other words, the action ANN ($\hat{W}^T \sigma(\hat{V}^T \hat{x})$) not only generates control signals to track the desired trajectory and cancel the unknown, uncertain, nonlinear dynamics but also minimises the cost function. At the same time, the critic agent ($0.5\hat{F}^T \sigma_f(\hat{x})$) estimates the cost function (29) and tunes the ANN weights. Clearly, through an adaptive learning and minimisation of the adaptive critic IT2FNN-based cost function in (29), an optimal or near optimal control action will be resulted. In fact, the proposed controller, that is designed in the sequel, not only stabilises the closed-loop system but also minimises the cost function simultaneously [52], [54].

Now, take the following control rule into account:

$$\begin{cases} \tau = \mathcal{D}^T \left(-K_p \frac{\ell_2}{k_o} - K_p \Lambda_\varepsilon \frac{\epsilon_e}{\sqrt{1 + \|\epsilon_e\|^2}} - \hat{W}^T \sigma(\hat{V}^T \hat{x}) + \hat{h} \right), \\ \hat{h} = -k_1 \mathcal{Q}_1 \hat{\varepsilon}_f - k_2 \mathcal{Q}_2 \hat{\varepsilon}_f - k_3 \mathcal{Q}_3 \hat{\varepsilon}_f - k_4 \mathcal{Q}_4 \hat{\varepsilon}_f - H \hat{p}, \end{cases} \quad (30)$$

where $K_p \in \mathbb{R}^{6 \times 6}$ is a gain matrix, $k_j, j = 1, \dots, 4$, are control parameters, $H := \text{diag}[\tanh(\hat{\varepsilon}_{f1}/c_{1r}), \dots, \tanh(\hat{\varepsilon}_{f6}/c_{6r})]$, \hat{p} is an upper bound estimation of some unknown parameters that are defined in the sequel, and the auxiliary terms are

$$\begin{cases} \mathcal{Q}_1 = (\|\sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}\| + \|\sigma(\hat{V}^T \hat{x})\|)^2 \\ \quad \times \|\sigma_f^T(\hat{x}) \hat{F}\|^2, \\ \mathcal{Q}_2 = \|\hat{x}\|^2 \|\sigma_f^T(\hat{x}) \hat{F} \hat{W}^T \sigma'(\hat{V}^T \hat{x})\|^2, \\ \mathcal{Q}_3 = \|\sigma_f(\hat{x}) \sigma'(\hat{V}^T \hat{x}) \hat{W}\|_F^2, \\ \mathcal{Q}_4 = (\|\sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}\| + \|\sigma(\hat{V}^T \hat{x})\|)^2 \\ \quad + \|\hat{x}\|^2 \|\hat{W}^T \sigma'(\hat{V}^T \hat{x})\|_F^2. \end{cases} \quad (31)$$

Then, the adaptive laws are designed here:

$$\begin{aligned}\dot{\hat{W}} &= \Gamma_w \left(\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x} \right) \\ &\quad \times \mathbf{1}, \left(\ell_2/k_o + \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2) \right)^{-\frac{1}{2}} \\ &\quad + 0.5 \|\ell_2/k_o + \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2) \right)^{-\frac{1}{2}} \hat{F}^T \sigma_f(\hat{x}) \Big)^T \\ &\quad - \delta_w \Gamma_w \hat{W},\end{aligned}\quad (32)$$

$$\begin{aligned}\dot{\hat{V}} &= \Gamma_v \hat{x} \left(\ell_2/k_o + \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2) \right)^{-\frac{1}{2}} \\ &\quad + 0.5 \|\ell_2/k_o + \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2) \right)^{-\frac{1}{2}} \hat{F}^T \sigma_f(\hat{x}) \Big)^T \\ &\quad \times \hat{W}^T \sigma'(\hat{V}^T \hat{x}) - \delta_v \Gamma_v \hat{V},\end{aligned}\quad (33)$$

$$\begin{aligned}\dot{\hat{F}} &= \Gamma_f \left\| \ell_2/k_o + \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2) \right\|^{-\frac{1}{2}} \left\| \sigma_f(\hat{x}) \right. \\ &\quad \times \left(\hat{W}^T \sigma(\hat{V}^T \hat{x}) \right)^T - \delta_f \Gamma_f \hat{F},\end{aligned}\quad (34)$$

$$\begin{aligned}\dot{\hat{p}} &= Q \left[H(\hat{\varepsilon}_f) (\ell_2/k_o + \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2) \right)^{-\frac{1}{2}} \\ &\quad - \Theta(\hat{p} - p^0) \Big],\end{aligned}\quad (35)$$

where $\Gamma_w \in \mathbb{R}^{(N_h+1) \times (N_h+1)}$, $\Gamma_v \in \mathbb{R}^{(N_i+1) \times (N_i+1)}$, $\Gamma_f \in \mathbb{R}^{2N \times 2N}$ and $Q, \Theta \in \mathbb{R}^{N_o \times N_o}$ are adaptive gains, $\delta_w, \delta_v, \delta_f \in \mathbb{R}^+$ are design parameters, and p^0 is a design vector. Replacing (30) into (16) and using (23) gives

$$\begin{aligned}M(\epsilon_e) \dot{\varepsilon}_f &= -C(\epsilon_e) \varepsilon_f - D(\epsilon_e) \varepsilon_f + \gamma - K_p \hat{\varepsilon}_f + \chi \\ &\quad + \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) - H \hat{p} \\ &\quad + \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \tilde{V}^T \hat{x} + r_t - k_1 \varrho_1 \hat{\varepsilon}_f \\ &\quad - k_2 \varrho_2 \hat{\varepsilon}_f - k_3 \varrho_3 \hat{\varepsilon}_f - k_4 \varrho_4 \hat{\varepsilon}_f,\end{aligned}\quad (36)$$

where $\chi = \delta_p + e_x(x) \in \mathbb{R}^{N_o}$ is bounded according to $|\chi_i| \leq p_i, i = 1, \dots, N_o$, where $p_i \in \mathbb{R}^+$ and $p = [p_1, \dots, p_{N_o}]^T$. Note that in the design procedure (for the example refer to (15)) and the proposed control law in (30), the saturation functions along with the compensation of the input saturation nonlinearity via the proposed hybrid RL method are utilized to handle the input constraint. Thus, the proposed controller generates input control signals with lowest possible amplitude. The schematic block diagram of the proposed control system is illustrated in Fig. 2. The equations of the system are given by (30)-(31) with the observer definition in (21) and (22) and adaptive laws (32)-(35).

A. STABILITY ANALYSIS

Theorem 1: Take into account the kinematic and dynamic of AUVs which are given by (2) and the transformed model (36). In accordance with Assumptions 1-3, the designed controller (30), adaptive rules (32)-(35) and state observer (21)-(22) with the reinforcement signal (29), ensure that all signals of the closed-loop system are bounded and the tracking errors are SGUUB which have a tendency toward the origin, and the pre-set output restrictions are never infringed.

Proof: See the Appendix section. \square

IV. SIMULATION STUDIES

In order to show the contributions of the proposed method, comparative simulations with two other controllers that are not RL-based controllers are performed which are the controller of Elhaki and Shojaei [6] and a non-prescribed performance output-feedback controller (NPPOFC) whose equations are given in [6]. Initial conditions of the AUVs are set as $q(0) = [10, 10, 10, 45\pi/180, 45\pi/180, 45\pi/180]^T$. The control parameters are as follows: $\eta_{10} = \eta_{20} = \eta_{30} = 2, \eta_{40} = \eta_{50} = \eta_{60} = 1, \eta_{1\infty} = \eta_{2\infty} = \eta_{3\infty} = 0.01, \eta_{4\infty} = \eta_{5\infty} = \eta_{6\infty} = 0.02, a_1 = a_2 = a_3 = 0.055, a_4 = a_5 = a_6 = 0.1, \alpha_1 = \alpha_2 = 5, \alpha_3 = 7, \alpha_4 = \alpha_5 = \alpha_6 = 1, \beta_1 = \beta_2 = 5, \beta_3 = 7, \beta_4 = \beta_5 = \beta_6 = 1.2, \Lambda_\varepsilon = I_6, k_o = \lambda_1 = 10, \bar{\sigma} = 1, \underline{\sigma} = 0.1, m_1 = -10, m_2 = -6, m_3 = -2, m_4 = 2, m_5 = 6, m_6 = 10, N = N_h = 6, \Gamma_w = 0.01 I_{N_h}, \delta_w = 1, \Gamma_v = 2 I_{N_i+1}, \delta_v = 0.1, \Gamma_f = I_{2N}, \delta_f = 0.01, Q = 2 \text{diag}[0.1, 0.1, 0.1, 1, 1, 1], \Theta = 0.1 \text{diag}[3, 1.5, 10, 1, 1, 1], p^0 = [1.7, 1.7, 1.7, 1.7, 1.7, 1.7]^T, c_{1r} = \dots = c_{6r} = 0.5, k_1 = \dots = k_4 = 1.5, K_p = 8 \text{diag}[1, 1, 1, 1, 1, 1].$

A. SCENARIO 1

In this scenario, Fig. 4 displays the results for the proposed method along with the comparisons. As it can be seen in Fig. 4(a), all control methods forced the AUVs to track the desired trajectory, but the AUV equipped with the proposed controller achieved a softer tracking than the controller in [6] and NPPOFC. Figs. 4(c)-4(h) show that the tracking errors of the proposed method and controller Elhaki and Shojaei [6] are easily kept inside the predefined funnel-shaped set without any concern about gain tuning. However, the tracking performance of NPPOFC seriously depends on the exact selection of control parameters and a random choice of the control parameters cannot ensure a good tracking performance. Moreover, the control signals that are shown in Figs. 4(i)-4(j) demonstrate a noticeable superiority for the proposed controller than the other two controllers due to the use of saturation functions in the design procedure and compensating the actuators saturation nonlinearity by the hybrid RL method, and the output patterns of the controller Elhaki and Shojaei [6] and NPPOFC are pretty fluctuated and saturated. This fact makes a compromise between the feasibility of the generated control signals by the proposed controller in Figs. 4(i)-4(j) and the convergence behaviour of the tracking errors in Figs. 4(c)-4(d). However, since the proposed control scheme is designed based on the PPC method, it never breaches the performance bounds. Moreover, in the proposed PPC control framework the transient performance could be improved by adjusting the PPC parameters. More simulation results on the improvement of the convergence behavior of the tracking errors are given in the next scenario. As reflected by the simulations, the amelioration in the outcomes are obviously notable which approve the use of the proposed novel NFRLC method in practice. Indeed, the strong estimation and compensation capability of the proposed estimation policy for uncertainties and nonlinearities,

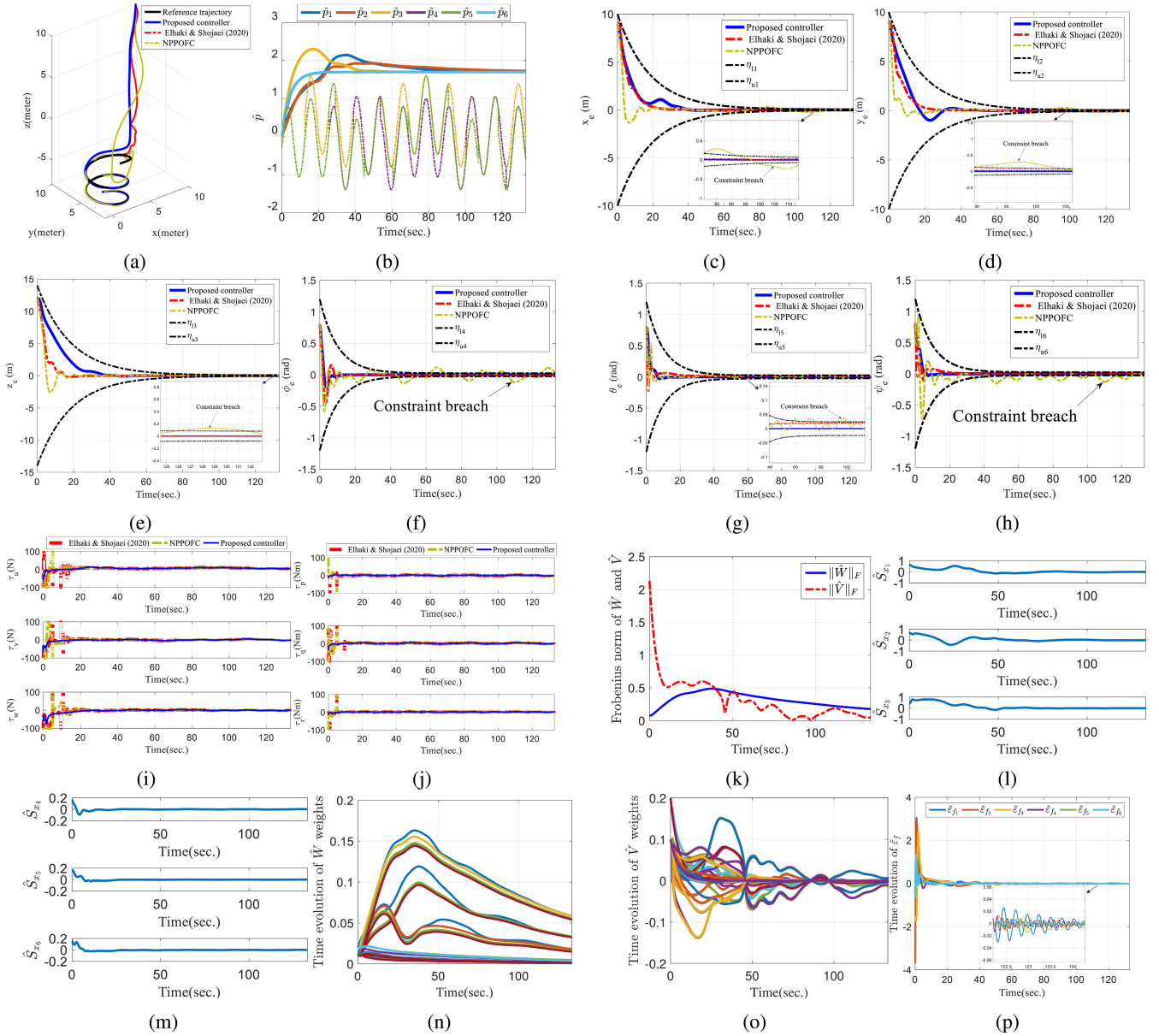


FIGURE 4. Simulation results for Scenario 1: (a) Trajectories of the AUVs in xyz space, (b) an upper bound approximation of unknown vector p ; comparison results for the tracking errors with performance bounds: (c) x_e , (d) y_e , (e) z_e , (f) ϕ_e , (g) θ_e , (h) ψ_e ; comparison results for the control signals: (i) input forces, (j) input torques, (k) the Frobenius norms of approximated NN weighting matrices, (l) linear reinforcement signals, (m) angular reinforcement signals; the growth of (n) \hat{W} and (o) \hat{V} weights, and (p) HGO estimation errors.

which is a result of the fusion between an adaptive MNN, adaptive IT2FNN, and the saturation functions, with a strong stability analysis led to generate substantially smooth and suitable output signals for the proposed controller in both transient and steady-state behaviours without the need for velocity measurements. Also, the upper bound of nonparametric uncertainties are estimated well in Fig. 4(b) and Frobenius norms of MNN matrices are depicted in Fig. 4(k). The RL signals are shown in Figs. 4(l)–4(m) and the time evolution of MNN weights are shown by Fig. 4(n) and Fig. 4(o). Fig. 4(p) presents the HGO estimation error, from which it can be realized that the designed HGO is able to present

estimations of the velocities adequately well. In this simulation, the CPU elapsed time for the proposed controller, the controller of Elhaki and Shojaei [6], and NPPOFC are 62, 46, and 15 ms, respectively, in a computer with a 2.4 GHz intel Core i5 processor. This result shows that the proposed hybrid RL controller has a little more computational cost with respect to the controller in [6] and NPPOFC at the cost of generating considerably better control action.

B. SCENARIO 2

In order to improve the trajectory tracking performance of the proposed controller by gain tuning, another simulation is

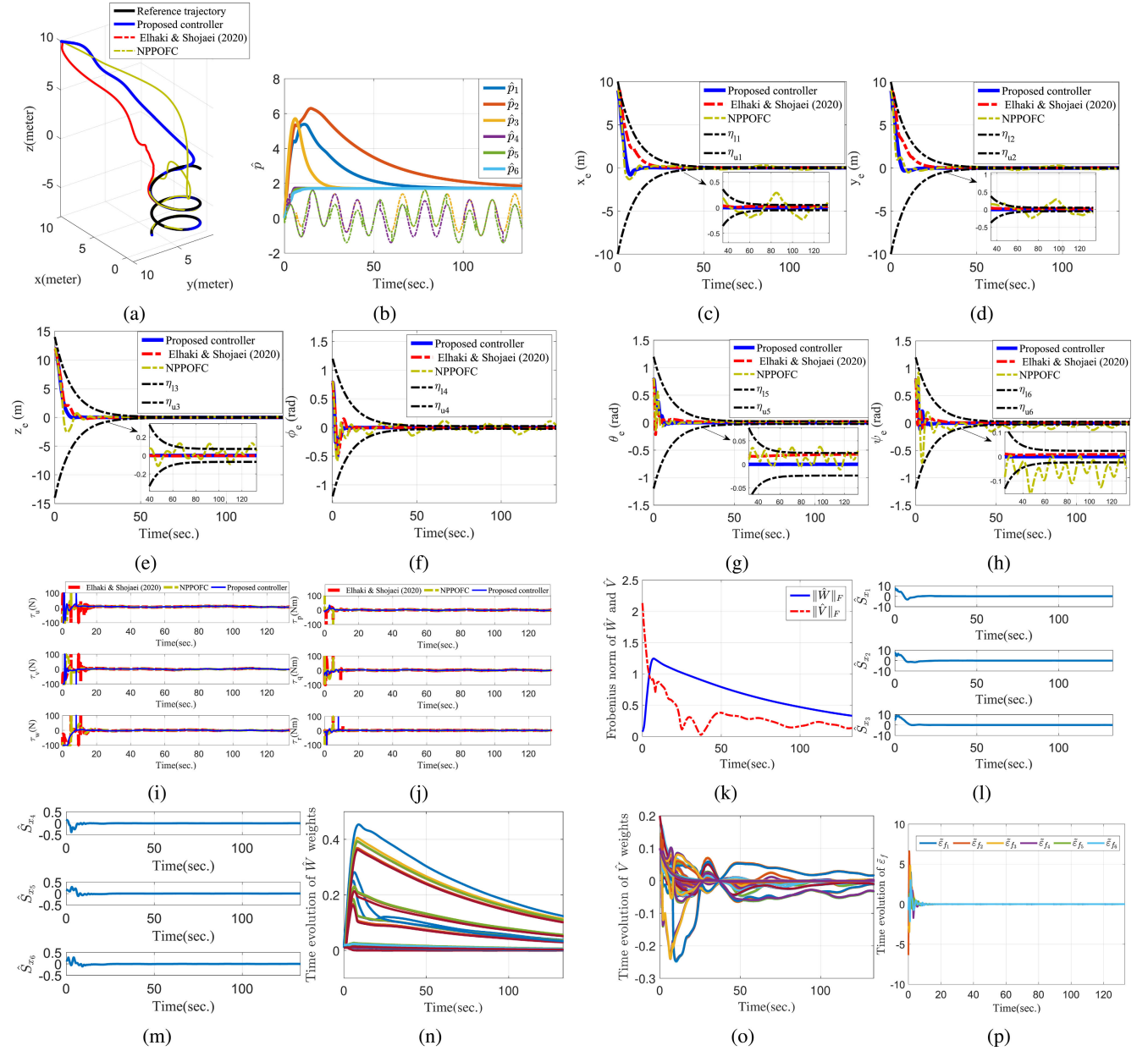


FIGURE 5. Simulation results for Scenario 2: (a) Trajectories of the AUVs in xyz space, (b) an upper bound approximation of unknown vector p ; comparison results for the tracking errors with performance bounds: (c) x_e , (d) y_e , (e) z_e , (f) ϕ_e , (g) θ_e , (h) ψ_e ; comparison results for the control signals: (i) input forces, (j) input torques, (k) the Frobenius norms of approximated NN weighting matrices, (l) linear reinforcement signals, (m) angular reinforcement signals; the growth of (n) \hat{W} and (o) \hat{V} weights, and (p) HGO estimation errors.

performed here. The adjusted control parameters are $a_1 = a_2 = a_3 = 0.1$, $\Lambda_\varepsilon = \text{diag}[12, 12, 12, 1, 1, 1]$ and other control parameters are similar to Scenario 1. Fig. 5(a) shows that the AUV equipped with the proposed controller converged considerably faster to the desired trajectory than other AUVs. As it can be seen from Figs. 5(c)-(e), the convergence behavior of the tracking errors significantly improved which is the result of this proper gain tuning. Hence, the advantage of the proposed controller in both transient and steady-state behaviors of the tracking errors is crystal-clear in Figs. 5(c)-(h). However, as shown by Figs. 5(i)-(j), this improvement of the

convergence for the proposed controller demands more control action than Scenario 1, but the generated control signals of the proposed controller are still more feasible than other two controllers. The other bounded signals of the suggested closed-loop system are also depicted in Figs. 5(b), (k)-(p). To put it in a nutshell, this scenario is considered to show that users may tune the convergence behavior of the closed-loop control system and can make a nice trade-off between the feasibility of the control action and convergence properties by a careful gain adjustment. More simulation results on the improvement of the convergence behavior of the tracking

errors with re-tuning of PPC parameters show the better performance of our controller. Nonetheless, they are omitted here and left to the interested readers.

V. CONCLUSION

In this paper, an intelligent saturated output feedback framework has been proposed for fully-actuated AUVs subjected to the unmodeled dynamics, actuator saturation and ocean disturbances. An adaptive robust controller has been used to reject exogenous disturbances. By designing an adaptive critic IT2FNN agent and an adaptive MNN, an adaptive hybrid NFRLC method was introduced to cope with dimensionality problem and to compensate every type of NLIP uncertainties and unmodeled dynamics. The suggested control model was independent of velocity sensors by means of a HGO, and the output constraints have been guaranteed via the PPC method. In conjunction with adaptive NFRLC and output feedback problems, a robust output feedback-based AC-NFRLC strategy was designed that did not engage with the system dynamics. In addition, by applying a saturation function into the design procedure along with the learning and compensating the actuators nonlinearity deeply, an effective saturated controller was developed. Moreover, by performing the Lyapunov stability analysis, it was proven that the tracking errors are SGUUB and all signals of the closed-loop control system were bounded and the tracking errors were also forced to evolve inside a funnel-shaped set to guarantee a prescribed performance. The performance of the proposed control structure could also be improved by employing more advanced control methods. For example, guaranteeing small overshoots and unknown initial errors are the subjects left for the future studies.

APPENDIX A

PROOF OF THEOREM 1

Proof: The following Lyapunov function is taken into account:

$$E = \sqrt{(1 + \epsilon_e^T \epsilon_e)} - 1 + 0.5 \epsilon_f^T M(\epsilon_e) \epsilon_f + 0.5 \tilde{p}^T Q^{-1} \tilde{p} + 0.5 \text{tr} \{ \tilde{W}^T \Gamma_w^{-1} \tilde{W} \} + 0.5 \text{tr} \{ \tilde{V}^T \Gamma_v^{-1} \tilde{V} \} + 0.5 \text{tr} \{ \tilde{F}^T \Gamma_f^{-1} \tilde{F} \}, \quad (37)$$

where $\tilde{W} = W^* - \hat{W}$, $\tilde{V} = V^* - \hat{V}$, $\tilde{F} = F^* - \hat{F}$, and $\tilde{p} = \hat{p} - p$. Eq. (37) is bounded by

$$\lambda_{x_\varepsilon} \|x_\varepsilon\|^2 \leq \lambda_l \|z_l\|^2 \leq E(t) \leq \lambda_u \|z_u\|^2, \quad (38)$$

where $x_\varepsilon = [\epsilon_e^T (1 + \|\epsilon_e\|^2)^{-\frac{1}{2}}, \epsilon_f^T]^T$ and

$$\begin{aligned} \lambda_{x_\varepsilon} &= 0.5 \min\{1, \lambda_m\}, \\ \lambda_l &= 0.5 \min\{1, \lambda_m, \lambda_{\min}\{\Gamma_w^{-1}\}, \lambda_{\min}\{\Gamma_v^{-1}\}, \\ &\quad \lambda_{\min}\{\Gamma_f^{-1}\}, \lambda_{\min}\{Q^{-1}\}\}, \\ \lambda_u &= 0.5 \max\{1, \lambda_M, \lambda_{\max}\{\Gamma_w^{-1}\}, \lambda_{\max}\{\Gamma_v^{-1}\}, \\ &\quad \lambda_{\max}\{\Gamma_f^{-1}\}, \lambda_{\max}\{Q^{-1}\}\}, \\ z_l &= [x_\varepsilon^T, \tilde{w}_{11}, \dots, \tilde{w}_{(N_h+1)N_o}, \tilde{v}_{11}, \dots, \tilde{v}_{(N_l+1)N_h}, \end{aligned}$$

$$\begin{aligned} &\tilde{f}_{11}, \dots, \tilde{f}_{2N \times m}, \tilde{p}^T]^T, \\ z_u &= [\epsilon_e^T, \epsilon_f^T, \tilde{w}_{11}, \dots, \tilde{w}_{(N_h+1)N_o}, \tilde{v}_{11}, \dots, \tilde{v}_{(N_l+1)N_h}, \\ &\quad \tilde{f}_{11}, \dots, \tilde{f}_{2N \times m}, \tilde{p}^T]^T. \end{aligned}$$

The derivative of (37) by using (36), item (i) of P1.3 and $\tilde{\epsilon}_f = \hat{\epsilon}_f - \epsilon_f$ gives

$$\begin{aligned} \dot{E} &= (1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \epsilon_e^T \dot{\epsilon}_e + \frac{1}{2} \epsilon_f^T \dot{M}(\epsilon_e) \epsilon_f + \epsilon_f^T M(\epsilon_e) \dot{\epsilon}_f \\ &\quad - \text{tr} \{ \tilde{W}^T \Gamma_w^{-1} \dot{\tilde{W}} \} - \text{tr} \{ \tilde{V}^T \Gamma_v^{-1} \dot{\tilde{V}} \} - \text{tr} \{ \tilde{F}^T \Gamma_f^{-1} \dot{\tilde{F}} \} \\ &\quad + \tilde{p}^T Q^{-1} \dot{\tilde{p}} \\ &= -(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \epsilon_e^T \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \\ &\quad + (1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \epsilon_e^T \epsilon_f - \epsilon_f^T D(\epsilon_e) \epsilon_f + \epsilon_f^T \gamma \\ &\quad + \epsilon_f^T (\chi - H\hat{p}) + \epsilon_f^T \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) \\ &\quad + \epsilon_f^T \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x} + \epsilon_f^T r_t - \epsilon_f^T \sum_{j=1}^4 k_j Q_j \hat{\epsilon}_f \\ &\quad - \text{tr} \{ \tilde{W}^T \Gamma_w^{-1} \dot{\tilde{W}} \} - \text{tr} \{ \tilde{V}^T \Gamma_v^{-1} \dot{\tilde{V}} \} - \text{tr} \{ \tilde{F}^T \Gamma_f^{-1} \dot{\tilde{F}} \} \\ &\quad + \tilde{p}^T Q^{-1} \dot{\tilde{p}} - \epsilon_f^T K_p \epsilon_f - \epsilon_f^T K_p \tilde{\epsilon}_f. \end{aligned} \quad (39)$$

Using update rules (32)-(35) yields

$$\begin{aligned} \dot{E} &= -(1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \epsilon_e^T \Lambda_\varepsilon \epsilon_e (1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \\ &\quad + (1 + \|\epsilon_e\|^2)^{-\frac{1}{2}} \epsilon_e^T \epsilon_f - \epsilon_f^T D(\epsilon_e) \epsilon_f + \epsilon_f^T \gamma \\ &\quad - \text{tr} \{ \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) \tilde{\epsilon}_f \} + \epsilon_f^T r_t \\ &\quad - \frac{1}{2} \text{tr} \{ \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) \|\hat{\epsilon}_f\| \sigma_f^T(\hat{x}) \hat{F} \} \\ &\quad - \text{tr} \{ \tilde{V}^T \hat{x} \tilde{\epsilon}_f^T \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \} - \epsilon_f^T K_p \epsilon_f - \epsilon_f^T K_p \tilde{\epsilon}_f \\ &\quad - \frac{1}{2} \text{tr} \{ \tilde{V}^T \hat{x} \|\hat{\epsilon}_f\| \sigma_f^T(\hat{x}) \hat{F} \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \} + \delta_v \text{tr} \{ \tilde{V}^T \hat{V} \} \\ &\quad - \text{tr} \{ \tilde{F}^T \|\hat{\epsilon}_f\| \sigma_f^T(\hat{x}) \sigma^T(\hat{V}^T \hat{x}) \hat{W} \} + \delta_w \text{tr} \{ \tilde{W}^T \hat{W} \} \\ &\quad + \delta_f \text{tr} \{ \tilde{F}^T \hat{F} \} - \epsilon_f^T \sum_{j=1}^4 k_j Q_j \hat{\epsilon}_f + \epsilon_f^T (\chi - H\hat{p}) \\ &\quad + \tilde{p}^T [H(\hat{\epsilon}_f) \hat{\epsilon}_f - \Theta(\hat{p} - p^0)]. \end{aligned} \quad (40)$$

By applying the next inequalities via using (24) and Lemma 1:

$$\begin{aligned} \|\epsilon_f^T r_t\| &\leq \frac{1}{2} \|W\|_F^2 + \frac{1}{2} \|V\|_F^2 + \frac{1}{2} \|\epsilon_f\|^2 Q_4, \\ &\quad - \text{tr} \{ \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) \tilde{\epsilon}_f \} \\ &\quad - \text{tr} \{ \tilde{V}^T \hat{x} \tilde{\epsilon}_f^T \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \} \leq \frac{1}{2} \|\tilde{W}\|_F^2 + \frac{1}{2} \|\tilde{V}\|_F^2 \\ &\quad + \frac{1}{2} B_\varepsilon^2 Q_4, \\ &\quad - \frac{1}{2} \text{tr} \{ \tilde{W}^T (\sigma(\hat{V}^T \hat{x}) - \sigma'(\hat{V}^T \hat{x}) \hat{V}^T \hat{x}) \|\hat{\epsilon}_f\| \sigma_f^T(\hat{x}) \hat{F} \} \\ &\leq \frac{1}{4} \|\tilde{W}\|_F^2 + \frac{1}{2} \|\epsilon_f\|^2 Q_1 + \frac{1}{2} B_\varepsilon^2 Q_1, \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2} \text{tr} \{ \tilde{V}^T \hat{x} \| \hat{\epsilon}_f \| \sigma_f^T(\hat{x}) \hat{F} \hat{W}^T \sigma'(\hat{V}^T \hat{x}) \} \leq \frac{1}{4} \| \tilde{V} \|_F^2 \\
& + \frac{1}{2} \| \epsilon_f \|^2 \varrho_2 + \frac{1}{2} B_\epsilon^2 \varrho_2, \\
& -\text{tr} \{ \tilde{F}^T \| \hat{\epsilon}_f \| \sigma_f^T(\hat{x}) \sigma^T(\hat{V}^T \hat{x}) \hat{W} \} \leq \| \tilde{F} \|_F^2 + \frac{1}{2} \| \epsilon_f \|^2 \varrho_3 \\
& + \frac{1}{2} B_\epsilon^2 \varrho_3, \\
& -\epsilon_f^T K_p \tilde{\epsilon}_f \leq \frac{1}{2} \lambda_{\max} \{ K_p \} \| \epsilon_f \|^2 + \frac{1}{2} \lambda_{\max} \{ K_p \} B_\epsilon^2, \\
& \epsilon_f^T \gamma \leq \frac{\iota_1}{2} \| \epsilon_f \|^2 + \frac{\iota_1}{2} \| x_\epsilon \|^2 + \frac{\iota_2}{2} \| \epsilon_f \|^2 + \frac{\iota_2}{2} \| x_\epsilon \|^4, \\
& (1 + \| \epsilon_e \|^2)^{-\frac{1}{2}} \epsilon_e^T \epsilon_f \leq \frac{1}{2} \| \epsilon_e \|^2 / (1 + \| \epsilon_e \|^2) + \frac{1}{2} \| \epsilon_f \|^2,
\end{aligned}$$

Eq. (40) is simplified as follows:

$$\begin{aligned}
\dot{E} \leq & -\beta_m \| x_\epsilon \|^2 + \epsilon_f^T (\chi - H\hat{p}) + \sum_{j=1}^4 \xi_j \varrho_j \\
& + \delta_f \text{tr} \{ \tilde{F}^T \hat{F} \} + \delta_v \text{tr} \{ \tilde{V}^T \hat{V} \} + \delta_w \text{tr} \{ \tilde{W}^T \hat{W} \} \\
& + \frac{1}{2} \| W \|_F^2 + \frac{1}{2} \| V \|_F^2 + \frac{3}{4} \| \tilde{W} \|_F^2 + \| \tilde{F} \|_F^2 \\
& + \frac{3}{4} \| \tilde{V} \|_F^2 + \tilde{p}^T [H(\hat{\epsilon}_f) \hat{\epsilon}_f - \Theta(\hat{p} - p^0)] \\
& + \frac{\iota_1}{2} \| x_\epsilon \|^2 + \frac{\iota_2}{2} \| x_\epsilon \|^4 + \frac{1}{2} \lambda_{\max} \{ K_p \} B_\epsilon^2, \quad (41)
\end{aligned}$$

where $\beta_m = \min \{ \lambda_{\min} \{ \Lambda_\epsilon \} - 0.5, \lambda_{\min} \{ K_p + D(\epsilon_e) \} - 0.5 \lambda_{\max} \{ K_p \} - \iota_1/2 - \iota_2/2 - 0.5 \}$ and $\xi_j \varrho_j = (- (0.5k_j - 0.5) \| \epsilon_f \|^2 + (0.5k_j + 0.5) B_\epsilon^2) \varrho_j$. So, if $\| \epsilon_f \| \geq B_\epsilon \sqrt{(k_j + 1)/(k_j - 1)}$ be true with $k_j > 1$, one has $\xi_j \varrho_j \leq 0$. By using the Young's Inequality [55] and the following relation [56]:

$$\begin{aligned}
& \tilde{p}^T [H(\hat{\epsilon}_f) \hat{\epsilon}_f - \Theta(\hat{p} - p^0)] + \epsilon_f^T (\chi - H\hat{p}) \\
& \leq 1.5B_\epsilon^2 \\
& + 0.2785[c_{1r}, \dots, c_{6r}]p + \| p \|^2 - \frac{1}{2} \lambda_{\min} \{ \Theta \} \| \tilde{p} \|^2 \\
& + 0.5 \| \tilde{p} \|^2 + 0.5(p - p^0)^T \Theta(p - p^0),
\end{aligned}$$

Eq. (41) is rewritten as

$$\begin{aligned}
\dot{E} \leq & -(\beta_m - 0.5\iota_1 - 0.5\iota_2 \| x_\epsilon \|^2) \| x_\epsilon \|^2 \\
& - (0.5\lambda_{\min} \{ \Theta \} - 0.5) \| \tilde{p} \|^2 - (\delta_w(1 - \frac{1}{2k^2}) - \frac{3}{4}) \| \tilde{W} \|_F^2 \\
& - (\delta_v(1 - \frac{1}{2k^2}) - \frac{3}{4}) \| \tilde{V} \|_F^2 - (\delta_f(1 - \frac{1}{2k^2}) - 1) \| \tilde{F} \|_F^2 \\
& + \Xi, \quad (42)
\end{aligned}$$

where $k > \sqrt{2}/2$, $\Xi = 0.2785[c_{1r}, \dots, c_{6r}]p + 1.5B_\epsilon^2 + \| p \|^2 + 0.5(p - p^0)^T \Theta(p - p^0) + 0.5\lambda_{\max} \{ K_p \} B_\epsilon^2 + 0.5\delta_w k^2 \| W \|_F^2 + 0.5\delta_v k^2 \| V \|_F^2 + 0.5\delta_f k^2 \| F \|_F^2 + 0.5 \| W \|_F^2 + 0.5 \| V \|_F^2$. If the following condition be true for β_m :

$$\beta_m > 0.5\iota_1 + 0.5\iota_2 \| x_\epsilon \|^2, \quad (43)$$

one has

$$\begin{aligned}
\dot{E} \leq & -c_1 \| x_\epsilon \|^2 - c_2 \| \tilde{W} \|_F^2 - c_3 \| \tilde{V} \|_F^2 - c_4 \| \tilde{F} \|_F^2 \\
& - c_5 \| \tilde{p} \|^2 + \Xi, \quad (44)
\end{aligned}$$

where $c_1 = \beta_m - 0.5\iota_1 - 0.5\iota_2 \| x_\epsilon \|^2$, $c_2 = \delta_w(1 - 1/2k^2) - 0.75$, $c_3 = \delta_v(1 - 1/2k^2) - 0.75$, $c_4 = \delta_f(1 - 1/2k^2) - 1$ and $c_5 = 0.5\lambda_{\min} \{ \Theta \} - 0.5$. Then, one has $\dot{E} \leq -c \| z_l \|^2 + \Xi$ where $c = \min \{ c_1, \dots, c_5 \}$. This means that \dot{E} is strictly negative when it is outside the compact set $\Omega_{z_l} = \{ z_l | 0 \leq \| z_l \| \leq \sqrt{\Xi/c} \} \forall t \geq 0$. Next, one may find that $E(t) \leq E(0) \leq \lambda_u \| z_u(0) \|^2$ by recalling (38). Consequently, one infers that $\| x_\epsilon \|^2 \leq \lambda_u / \lambda_{x_\epsilon} \| z_u(0) \|^2$. As a result, an adequate situation for (43) is $2\beta_m > \iota_1 + \iota_2(\lambda_u / \lambda_{x_\epsilon}) \| z_u(0) \|^2$ which means that the following region of attraction

$$R_A = \left\{ z_u \in \mathbb{R}^{18+\xi_n} | \| z_u \|^2 < \frac{\lambda_{x_\epsilon}(2\beta_m - \iota_1)}{\iota_2 \lambda_u} \right\}, \quad (45)$$

where $\xi_n = (N_h + 1)N_o + (N_i + 1)N_h + 2N \times m$, can include every initial condition by choosing control gains properly. By recalling (15), one deduces $\epsilon_e \in \mathcal{L}_\infty$. This situation implies that tracking errors, NN and fuzzy weights as well as variables estimation errors are SGUUB. Besides, by employing Assumptions 2-3 and $\| \tilde{\epsilon}_f \| \leq B_\epsilon$, one concludes that $\epsilon_e, \hat{\epsilon}_f, \hat{w}_{11}, \dots, \hat{w}_{(N_h+1)N_o}, \hat{v}_{11}, \dots, \hat{v}_{(N_i+1)N_h}, \hat{f}_{11}, \dots, \hat{f}_{2N \times m}, \hat{p} \in \mathcal{L}_\infty$. Because $\epsilon_e \in \mathcal{L}_\infty$ and Assumption 1 is considered, ϵ_e tends to a vicinity of the origin, then, the tracking errors converge to the origin vicinity guaranteeing output constraints and the boundaries conditions on the tracking errors, i.e. $\eta_{li}(t) \leq e_i(t) \leq \eta_{ui}(t)$ are never breached as $t \rightarrow \infty$, and the proof is accomplished here. \square

REFERENCES

- [1] T. Burrell, A. Montazeri, S. Monk, and C. J. Taylor, "Feedback control-based inverse kinematics solvers for a nuclear decommissioning robot," *IFAC-PapersOnLine*, vol. 49, no. 21, pp. 177–184, 2016.
- [2] M. Bandala, C. West, S. Monk, A. Montazeri, and C. J. Taylor, "Vision-based assisted tele-operation of a dual-arm hydraulically actuated robot for pipe cutting and grasping in nuclear environments," *Robotics*, vol. 8, no. 2, p. 42, Jun. 2019.
- [3] T. Burrell, C. West, S. D. Monk, A. Montazeri, and C. J. Taylor, "Towards a cooperative robotic system for autonomous pipe cutting in nuclear decommissioning," in *Proc. UKACC 12th Int. Conf. Control (CONTROL)*, Sep. 2018, pp. 283–288.
- [4] H. Nemati and A. Montazeri, "Analysis and design of a multi-channel time-varying sliding mode controller and its application in unmanned aerial vehicles," *IFAC-PapersOnLine*, vol. 51, no. 22, pp. 244–249, 2018.
- [5] S. Pyne, W. Jones, I. Crabbe, and E. Carr, "Submersible ROVs pioneer decommissioning of a legacy nuclear fuel storage pond-15368," *WM Symposia*, Tempe, AZ, USA, Tech. Rep. INIS-US-19-WM-15368, 2015.
- [6] O. Elhaki and K. Shojaei, "A robust neural network approximation-based prescribed performance output-feedback controller for autonomous underwater vehicles with actuators saturation," *Eng. Appl. Artif. Intell.*, vol. 88, Feb. 2020, Art. no. 103382.
- [7] Y.-H. Chang and W.-S. Chan, "Adaptive dynamic surface control for uncertain nonlinear systems with interval type-2 fuzzy neural networks," *IEEE Trans. Cybern.*, vol. 44, no. 2, pp. 293–304, Feb. 2014.
- [8] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.
- [9] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—I," *Inf. Sci.*, vol. 8, no. 3, pp. 199–249, Jan. 1975.
- [10] Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: Theory and design," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 5, pp. 535–550, Oct. 2000.
- [11] W. Peng, C. Li, G. Zhang, and J. Yi, "Interval type-2 fuzzy logic based transmission power allocation strategy for lifetime maximization of WSNs," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103269.
- [12] Q. Shi, H.-K. Lam, C. Xuan, and M. Chen, "Adaptive neuro-fuzzy PID controller based on twin delayed deep deterministic policy gradient algorithm," *Neurocomputing*, vol. 402, pp. 183–194, Aug. 2020.

- [13] P. M. Kebria, A. Khosravi, S. Nahavandi, D. Wu, and F. Bello, "Adaptive type-2 fuzzy neural-network control for teleoperation systems with delay and uncertainties," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 10, pp. 2543–2554, Oct. 2020.
- [14] J. Zhao, Y. Liu, L. Wang, and W. Wang, "A generalized heterogeneous type-2 fuzzy classifier and its industrial application," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 10, pp. 2287–2301, Oct. 2020.
- [15] X. Bu and Q. Qi, "Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design," *IEEE Trans. Fuzzy Syst.*, early access, Nov. 11, 2020, doi: [10.1109/TFUZZ.2020.3036706](https://doi.org/10.1109/TFUZZ.2020.3036706).
- [16] S. Luo, F. L. Lewis, Y. Song, and H. M. Ouakad, "Accelerated adaptive fuzzy optimal control of three coupled fractional-order chaotic electromechanical transducers," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 7, pp. 1701–1714, Jul. 2021, doi: [10.1109/TFUZZ.2020.2984998](https://doi.org/10.1109/TFUZZ.2020.2984998).
- [17] X. Min, Y. Li, and S. Tong, "Adaptive fuzzy output feedback inverse optimal control for vehicle active suspension systems," *Neurocomputing*, vol. 403, pp. 257–267, Aug. 2020.
- [18] S. Yang, J. Han, L. Xia, and Y.-H. Chen, "An optimal fuzzy-theoretic setting of adaptive robust control design for a lower limb exoskeleton robot system," *Mech. Syst. Signal Process.*, vol. 141, Jul. 2020, Art. no. 106706.
- [19] Y. Fan and Y. Li, "Adaptive fuzzy finite-time optimal control for switched nonlinear systems," *Optim. Control Appl. Methods*, vol. 41, no. 5, pp. 1616–1631, Sep. 2020.
- [20] E. Susanto, A. S. Wibowo, and E. G. Rachman, "Fuzzy swing up control and optimal state feedback stabilization for self-erecting inverted pendulum," *IEEE Access*, vol. 8, pp. 6496–6504, 2020.
- [21] E. Soares, P. P. Angelov, B. Costa, M. P. G. Castro, S. Nagesh Rao, and D. Filev, "Explaining deep learning models through rule-based approximation and visualization," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 8, pp. 2399–2407, Aug. 2021, doi: [10.1109/TFUZZ.2020.2999776](https://doi.org/10.1109/TFUZZ.2020.2999776).
- [22] Z. Sui, Z. Pu, J. Yi, and S. Wu, "Formation control with collision avoidance through deep reinforcement learning using model-guided demonstration," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 6, pp. 2358–2372, Jun. 2021, doi: [10.1109/TNNLS.2020.3004893](https://doi.org/10.1109/TNNLS.2020.3004893).
- [23] A. A. Khater, A. M. El-Nagar, M. El-Bardini, and N. El-Rabaie, "A novel structure of actor-critic learning based on an interval type-2 TSK fuzzy neural network," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 11, pp. 3047–3061, Nov. 2020.
- [24] W. He, H. Gao, C. Zhou, C. Yang, and Z. Li, "Reinforcement learning control of a flexible two-link manipulator: An experimental investigation," *IEEE Trans. Syst., Man, Cybern. Syst.*, early access, Mar. 5, 2020, doi: [10.1109/TSMC.2020.2975232](https://doi.org/10.1109/TSMC.2020.2975232).
- [25] S. Xue, B. Luo, and D. Liu, "Integral reinforcement learning based event-triggered control with input saturation," *Neural Netw.*, vol. 131, pp. 144–153, Nov. 2020.
- [26] M. Gheisarnajad and M. H. Khooban, "An intelligent non-integer PID controller-based deep reinforcement learning: Implementation and experimental results," *IEEE Trans. Ind. Electron.*, vol. 68, no. 4, pp. 3609–3618, Apr. 2021, doi: [10.1109/TIE.2020.2979561](https://doi.org/10.1109/TIE.2020.2979561).
- [27] G. A. P. de Moraes, L. B. Marcos, J. N. A. D. Bueno, N. F. de Resende, M. H. Terra, and V. Grassi, Jr., "Vision-based robust control framework based on deep reinforcement learning applied to autonomous ground vehicles," *Control Eng. Pract.*, vol. 104, Nov. 2020, Art. no. 104630.
- [28] X. Bu, Y. Xiao, and H. Lei, "An adaptive critic design-based fuzzy neural controller for hypersonic vehicles: Predefined behavioral nonaffine control," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 4, pp. 1871–1881, Aug. 2019.
- [29] D. Zhao, Z. Wang, G. Wei, and Q.-L. Han, "A dynamic event-triggered approach to observer-based PID security control subject to deception attacks," *Automatica*, vol. 120, Oct. 2020, Art. no. 109128.
- [30] J. Wang, Z. Zhou, C. Wang, and Z. Ding, "Cascade structure predictive observer design for consensus control with applications to UAVs formation flying," *Automatica*, vol. 121, Nov. 2020, Art. no. 109200.
- [31] Y. Zhang, J. Jin, and L. Huang, "Model-free predictive current control of PMSM drives based on extended state observer using ultralocal model," *IEEE Trans. Ind. Electron.*, vol. 68, no. 2, pp. 993–1003, Feb. 2021.
- [32] C. Kitsos, G. Besançon, and C. Prieur, "High-gain observer design for some semilinear reaction-diffusion systems: A transformation-based approach," *IEEE Control Syst. Lett.*, vol. 5, no. 2, pp. 629–634, Apr. 2021.
- [33] C. P. Bechlioulis and G. A. Rovithakis, "Prescribed performance adaptive control for multi-input multi-output affine in the control nonlinear systems," *IEEE Trans. Autom. Control*, vol. 55, no. 5, pp. 1220–1226, May 2010.
- [34] O. Elhaki and K. Shojaei, "Neural network-based target tracking control of underactuated autonomous underwater vehicles with a prescribed performance," *Ocean Eng.*, vol. 167, pp. 239–256, Nov. 2018.
- [35] O. Elhaki and K. Shojaei, "Observer-based neural adaptive control of a platoon of autonomous tractor-trailer vehicles with uncertain dynamics," *IET Control Theory Appl.*, vol. 14, no. 14, pp. 1898–1911, Sep. 2020.
- [36] O. Elhaki and K. Shojaei, "Robust prescribed performance-based control of autonomous tractor-trailers convoy with limited communication range," *Int. J. Syst. Sci.*, vol. 52, no. 3, pp. 555–582, Feb. 2021.
- [37] O. Elhaki and K. Shojaei, "Trajectory tracking control of underactuated autonomous marine surface vessels with guaranteed prescribed performance," in *Proc. 6th RSI Int. Conf. Robot. Mechatronics (ICRoM)*, Oct. 2018, pp. 22–27.
- [38] N.-N. Zhao, X.-Y. Ouyang, L.-B. Wu, and F.-R. Shi, "Event-triggered adaptive prescribed performance control of uncertain nonlinear systems with unknown control directions," *ISA Trans.*, vol. 108, pp. 121–130, Feb. 2021.
- [39] M. Hou and Y. Wang, "Data-driven adaptive terminal sliding mode control with prescribed performance," *Asian J. Control*, vol. 23, no. 2, pp. 774–785, Mar. 2021.
- [40] X. Shao, H. Si, and W. Zhang, "Fuzzy wavelet neural control with improved prescribed performance for MEMS gyroscope subject to input quantization," *Fuzzy Sets Syst.*, vol. 411, pp. 136–154, May 2021.
- [41] Y. Wang, B. Niu, H. Wang, N. Alotaibi, and E. Abozinadah, "Neural network-based adaptive tracking control for switched nonlinear systems with prescribed performance: An average dwell time switching approach," *Neurocomputing*, vol. 435, pp. 295–306, May 2021.
- [42] T. Zhou, C. Liu, X. Liu, H. Wang, and Y. Zhou, "Finite-time prescribed performance adaptive fuzzy control for unknown nonlinear systems," *Fuzzy Sets Syst.*, vol. 402, pp. 16–34, Jan. 2021.
- [43] K. D. Do and J. Pan, *Control of Ships and Underwater Vehicles: Design for Underactuated and Nonlinear Marine Systems*. London, U.K.: Springer, 2009.
- [44] M. Chen, "Disturbance attenuation tracking control for wheeled mobile robots with skidding and slipping," *IEEE Trans. Ind. Electron.*, vol. 64, no. 4, pp. 3359–3368, Apr. 2017.
- [45] J. Li, J. Li, Z. Wu, and Y. Liu, "Practical tracking control with prescribed transient performance for Euler-Lagrange equation," *J. Franklin Inst.*, vol. 357, no. 10, pp. 5809–5830, Jul. 2020.
- [46] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang, *Stable Adaptive Neural Network Control*, vol. 13. New York, NY, USA: Springer, 2013.
- [47] S. S. Ge and J. Zhang, "Neural-network control of nonaffine nonlinear system with zero dynamics by state and output feedback," *IEEE Trans. Neural Netw.*, vol. 14, no. 4, pp. 900–918, Jul. 2003.
- [48] K. P. Tee, S. S. Ge, and F. E. H. Tay, "Adaptive neural network control for helicopters in vertical flight," *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 4, pp. 753–762, Jul. 2008.
- [49] A. Mohammadzadeh and S. Ghaemi, "A modified sliding mode approach for synchronization of fractional-order chaotic/hyperchaotic systems by using new self-structuring hierarchical type-2 fuzzy neural network," *Neurocomputing*, vol. 191, pp. 200–213, May 2016.
- [50] A. Mohammadzadeh and T. Kumbasar, "A new fractional-order general type-2 fuzzy predictive control system and its application for glucose level regulation," *Appl. Soft Comput.*, vol. 91, Jun. 2020, Art. no. 106241.
- [51] C.-K. Lin, "H ∞ reinforcement learning control of robot manipulators using fuzzy wavelet networks," *Fuzzy Sets Syst.*, vol. 160, no. 12, pp. 1765–1786, Jun. 2009.
- [52] Q. Yang and S. Jagannathan, "Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 377–390, Apr. 2012.
- [53] Y. Luo, Q. Sun, H. Zhang, and L. Cui, "Adaptive critic design-based robust neural network control for nonlinear distributed parameter systems with unknown dynamics," *Neurocomputing*, vol. 148, pp. 200–208, Jan. 2015.
- [54] L. Tang, Y.-J. Liu, and S. Tong, "Adaptive neural control using reinforcement learning for a class of robot manipulator," *Neural Comput. Appl.*, vol. 25, no. 1, pp. 135–141, Jul. 2014.
- [55] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
- [56] J. Du, X. Hu, H. Liu, and C. L. P. Chen, "Adaptive robust output feedback control for a marine dynamic positioning system based on a high-gain observer," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 11, pp. 2775–2786, Nov. 2015.



nonlinear systems and Euler-Lagrange systems and robotics.

OMID ELHAKI was born in Esfahan, Iran, in May 1991. He received the B.Sc. degree in electrical engineering from Iran University of Applied Sciences and Technology (UAST), in 2015, and the M.Sc. degree (Hons.) in control engineering from Islamic Azad University, Najafabad Branch, Iran, in 2017. His research interests include adaptive and robust control, reinforcement learning, fuzzy control, and prescribed performance control with applications in uncertain autonomous



observer design, and non-linear control of discrete time systems, with a focus on applications in robotics.

DECLAN SHANAHAN was born in Macclesfield, Cheshire, U.K., in 1994. He received the B.E. degree in engineering from the University of Liverpool, Merseyside, U.K., in 2017, and the M.S. degree in mechanical engineering from Lancaster University, Lancashire, U.K., in 2019. He is currently a Research Associate with the Department of Engineering, Lancaster University. His research interests include dynamic modeling and simulation of hydraulic systems, sliding mode control,



observer design, and non-linear control of discrete time systems, with a focus on applications in robotics.

ALLAHYAR MONTAZERI received the B.Sc. degree in electrical engineering from Tehran University, Tehran, Iran, in 2000, and the M.Sc. and Ph.D. degrees (Hons.) in electrical engineering from Iran University of Science and Technology, Tehran, in 2002 and 2009, respectively. From 2010 to 2011, he joined Fraunhofer Institute, Germany, as a Research Fellow, and then carried on his research with Fraunhofer IDMT and Control Engineering Group, Ilmenau University of Technology, Germany, from 2011 to 2013. He has been a Visiting Research Scholar with the Control Engineering Group, ETH Zurich, Switzerland, and the Chemical Engineering Group, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. Since 2013, he has been an Assistant Professor with the Department of Engineering, Lancaster University, U.K. His research interests include wide range of areas on control theory and digital signal processing. He is particularly interested in adaptive signal processing and control, robust control, linear and nonlinear system identification, estimation theory, and evolutionary computing and optimization with applications in active noise and vibration control systems, and robotics. He is a fellow of the Higher Education Academy. He is currently serving on IFAC Technical Committees in adaptive and learning systems and modeling, identification, and signal processing. He was a recipient of the European Research Consortium on Informatics and Mathematics (ERCIM) and Humboldt research awards, in 2010 and 2011, respectively. His research was funded by different councils and industries in U.K., such as the Engineering and Physical Research Council, Sellafield Ltd., the National Nuclear laboratory, and Nuclear Decommissioning Authority.

...



KHOSHNAM SHOJAEI was born in Esfahan, Iran, in March 1981. He received the B.Sc., M.Sc., and Ph.D. degrees (Hons.) in electrical engineering from Iran University of Science and Technology (IUST), in 2004, 2007, and 2011, respectively. He is currently an Associate Professor with Najafabad Branch, Islamic Azad University. His research interests include adaptive control of nonlinear systems, control of autonomous robots including land, air and ocean vehicles, and navigation of mobile robots.